

# Fei Xue

## Employment

- 2021–current **Assistant Professor**, *Purdue University, USA*  
Spring 2025 **Visiting Research Associate**, *University of Chicago (IMSI), USA*  
2019–2021 **Postdoc Researcher**, *The University of Pennsylvania, USA*  
Supervisor: Hongzhe Li

## Education

- 2014–2019 **Ph.D. in Statistics**, *The University of Illinois at Urbana–Champaign, USA*  
Advisor: Annie Qu  
2010–2014 **B.S. in Mathematics and Applied Mathematics**, *Fudan University, China*

## Research Interests

Data integration, mobile health, missing data, mediation analysis, high-dimensional inference, and statistical genetics.

## Papers

1. **Fei Xue**, Bingxin Zhao (2025). High-dimensional statistical inference for linkage disequilibrium score regression and its cross-ancestry extensions. **Annals of Statistics**, 53(5), 1913–1937. <https://doi.org/10.1214/25-AOS2523>
2. **Fei Xue**, Annie Qu (2021). Integrating multi-source block-wise missing data in model selection. **Journal of the American Statistical Association**, 116:536, 1914–1927. <https://doi.org/10.1080/01621459.2020.1751176>
3. **Fei Xue**, Yanqing Zhang, Wenzhuo Zhou, Haoda Fu, Annie Qu (2022). Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data. **Journal of the American Statistical Association**, 117:539, 1438–1451. <https://doi.org/10.1080/01621459.2020.1862671>
4. **Fei Xue**, Xiwei Tang, Grace Kim, Allison Aiello, Karestan Koenen, Sandro Galea, Derek Wildman, Monica Uddin, Annie Qu (2022). Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort. **Journal of the American Statistical Association**, 117:540, 1669–1683. <https://doi.org/10.1080/01621459.2022.2089572>

5. Xiwei Tang, **Fei Xue**, Annie Qu (2021). Individualized multi-directional variable selection. **Journal of the American Statistical Association**, 116:535, 1280-1296. <https://doi.org/10.1080/01621459.2019.1705308>
6. Jiuchen Zhang, **Fei Xue**, Qi Xu, Jung-Ah Lee, Annie Qu (2024). Individualized dynamic latent factor model with application in mobile health data. **Biometrika**, 111(4), 1257-1275. <https://doi.org/10.1093/biomet/asae015>
7. **Fei Xue**, Annie Qu (2022). Semi-standard partial covariance variable selection when irrepresentable conditions fail. **Statistica Sinica**, 32: 1881-1909. <https://doi.org/10.5705/ss.202020.0495>
8. **Fei Xue**, Rong Ma, Hongzhe Li. Statistical inference for high-dimensional linear regression with blockwise missing data. **Statistica Sinica**, accepted. [https://www3.stat.sinica.edu.tw/ss\\_newpaper/SS-2022-0104\\_na.pdf](https://www3.stat.sinica.edu.tw/ss_newpaper/SS-2022-0104_na.pdf)
9. **Fei Xue**, Hongzhe Li. An empirical Bayes regression for multi-tissue eQTL data analysis. **Statistica Sinica**, accepted. [https://www3.stat.sinica.edu.tw/ss\\_newpaper/SS-2023-0201\\_na.pdf](https://www3.stat.sinica.edu.tw/ss_newpaper/SS-2023-0201_na.pdf)
10. Kecheng Wei, **Fei Xue**, Qi Xu, Yubai Yuan, Yuxia Zhang, Guoyou Qin, Agaz Wani, Allison Aiello, Derek Wildman, Monica Uddin, Annie Qu. Time-varying mediation analysis for incomplete data with application to DNA methylation study for PTSD. **Annals of Applied Statistics**, accepted. <https://www.biorxiv.org/content/10.1101/2024.02.06.579228v1>
11. Simona Buetti, **Fei Xue**, Qiwen Liu, Juyoen Hur, Gavin Jun Peng Ng, Wendy Heller (2020). Perceived control in the lab and in daily life impact emotion-induced temporal distortions. **Timing & Time Perception**, 9(1): 88-122. <https://doi.org/10.1163/22134468-bja10018>
12. Grace S Kim, Alicia Smith, **Fei Xue**, Vasiliki Michopoulos, Adriana Lori, Don L. Armstrong, Allison E. Aiello, Karestan C. Koenen, Sandro Galea, Derek E. Wildman, Monica Uddin (2019). Methylomic profiles reveal sex-specific shifts in leukocyte composition associated with post-traumatic stress disorder. **Brain, Behavior, and Immunity**, 81: 280–291. <https://doi.org/10.1016/j.bbi.2019.06.025>
13. Agaz Wani, Allison Aiello, Grace S Kim, **Fei Xue**, Andrew Ratanatharathorn, Annie Qu, Karestan Koenen, Sandro Galea, Derek Wildman, Monica Uddin (2021). The impact of psychopathology, social adversity and stress-relevant DNAm on prospective risk for post-traumatic stress: A machine learning approach. **Journal of Affective Disorders**, 282: 894-905. <https://www.biorxiv.org/content/10.1101/2020.09.25.313635v1>
14. Yujue Li, **Fei Xue**, Bingxuan Li, Yilin Yang, Zirui Fan, Juan Shu, Xiyao Wang, Jinjie Lin, Carlos Copana, Bingxin Zhao. Analyzing bivariate cross-trait genetic architecture in GWAS summary statistics with the BIGA cloud computing platform. Submitted. <https://doi.org/10.1101/2023.04.28.538585>
15. Huiming Xie, **Fei Xue**, Xiao Wang. Identifiable deep latent variable models for MNAR data. Under revision.

16. Xiaojing Sun, Bingxin Zhao, **Fei Xue**. Generalized heterogeneous functional model with applications to large-scale mobile health data. Submitted.
17. Jin Jin, Bingxuan Li, Xiyao Wang, Xiaochen Yang, Yujue Li, Ruofan Wang, Chenglong Ye, Juan Shu, Zirui Fan, **Fei Xue**, Tian Ge, Marylyn D. Ritchie, Bogdan Pasaniuc, Genevieve Wojcik, Bingxin Zhao. PennPRS: a centralized cloud computing platform for efficient polygenic risk score training in precision medicine. Submitted.
18. Antik Chakraborty, **Fei Xue**. Empirical Bayes data integration for multi-response regression. Under revision.
19. Rebecca A Deek, **Fei Xue**, James D Lewis, Hongzhe Li. Estimation and false discovery control of multivariate causal effects in omics studies. Submitted.
20. Huiming Xie, **Fei Xue**, Xiao Wang. Transition dynamics modeling of pre-trained accelerometry representations for time to diagnosis of Parkinson's disease. Submitted.
21. Yuexia Zhang, Yubai Yuan, **Fei Xue**, Kecheng Wei, Jin Zhou, Annie Qu. Causal mediation analysis of the effect of depression on Alzheimer's disease risk in older adults. Submitted.

## Book Chapters

1. Huiming Xie, **Fei Xue**, Xiao Wang. Generative Models for Missing Data in Applications of Generative AI , Springer, 2024. [https://link.springer.com/chapter/10.1007/978-3-031-46238-2\\_27](https://link.springer.com/chapter/10.1007/978-3-031-46238-2_27)

## Grants

- 2022-2026 **Principal Investigator**, *Integrative approaches with applications in eQTL analysis and randomized trials*, NSF DMS 2210860
- 2023-2026 **Co-Principal Investigator**, *Machine Learning for Predictive Breeding Technology*, Beck's Superior Hybrids, INC
- 2025-2026 **Co-Principal Investigator**, *Development of Advanced Analytical Tools Integrating External Datasets with MoTrPAC Data: Utilizing Statistical, Machine Learning, and AI Approaches to Uncover Molecular Mechanisms of Physical Activity*, NIH 1R03OD039853
- 2025-2026 **Co-Principal Investigator**, *Leveraging AI to Unlock the Potential of Emerging Data for Accelerated Underwriting: A General Framework and Case Study in Wearable-Based Mortality Analytics*, Society of Actuaries (SOA)

## Selected Awards

- 2024 **Seed for Success Acorn Award**, *Purdue University*, USA
- 2022 **IMS New Researchers Travel Award**, *Institute of Mathematical Statistics*, USA
- 2019 **ICSA New Researcher Award**, *International Chinese Statistical Association*, USA
- 2019 **IMS Hannan Graduate Student Travel Award**, *Institute of Mathematical Statistics*, USA

- 2018 **ASA Student Paper Award**, *Statistical Learning and Data Science Section of the American Statistical Association*, USA
- 2018 **Horace W. Norton Prize for Outstanding Thesis Research Finalist**, *Department of Statistics, UIUC*, USA
- 2018 **The Third Place in Student Poster Competition**, *The Fourth International Workshop on the Statistical Analyses of Multi-Outcome Data*, USA
- 2013 **Meritorious Winner**, *Mathematical Contest in Modeling*, USA
- 2010 **Bronze medal**, *Chinese Mathematical Olympiad*, China

## Students

Huiming Xie, Xiaochen Yang, Xiaojing Sun, Hyunwoo Chung, Yujue Li, Huayu Tang, Ziqi Dong, Huei-Yuan Yao

## Teaching Experience

### Instructor at Purdue

- Fall 2025 Stat 417: **Statistical Theory** (two sections)
- Fall 2024 Stat 417: **Statistical Theory** (two sections)
- Spring 2024 Stat 417: **Statistical Theory** (two sections)
- Spring 2023 Stat 528: **Introduction to Mathematical Statistics**
- Fall 2022 Stat 517: **Statistical Inference**
- Spring 2022 Stat 517: **Statistical Inference**
- Fall 2021 Stat 517: **Statistical Inference**

### Instructor at UIUC

- Fall 2016 Stat 200: **Statistical Analysis**  
Independently led a class of 40 undergraduate students, gave lectures, held office hours and designed exams and homework assignments

## Talks

### Invited

- Oct. 2025 **Seminar, Binghamton University, Binghamton, NY**  
Statistical Methods for Mobile Health Data
- July 2025 **NSF@75: Advancing Statistical Science for a Data-Driven World, online**  
An Empirical Bayes Regression for Multi-tissue eQTL Data Analysis
- April 2025 **Seminar, University of Texas at Dallas, Richardson, TX**  
Statistical Methods for Mobile Health Data
- April 2025 **IMSI Research Presentation, Chicago, IL**  
Generalized heterogeneous functional model with applications to mobile health data
- March 2025 **ENAR 2025 Spring Meeting, New Orleans, LA**  
Time-Varying Mediation Analysis for Incomplete Data with Application to DNA Methylation Study for PTSD
- March 2025 **Seminar, Rice University, Huston, TX**  
Statistical Methods for Mobile Health Data

- Feb. 2025 **Seminar, University of Louisville, Louisville, KY**  
 Statistical Methods for Mobile Health Data
- Nov. 2024 **Undergraduate Women in Science Program (WISP), West Lafayette, IN**  
 From data to discovery
- Nov. 2024 **SLDS 2024, Newport Beach, CA**  
 Generalized heterogeneous functional model with applications to mobile health data
- Sep. 2024 **Seminar, University of Kentucky, Lexington, KY**  
 Statistical Methods for Mobile Health Data
- Aug. 2024 **Conference on Innovative Statistics and Machine Learning for Data Science and AI (IMSL), Bend, OR**  
 Generalized heterogeneous functional model with applications to mobile health data
- July 2024 **EcoSta 2024 (hybrid), Beijing, China**  
 High-dimensional statistical inference for linkage disequilibrium score regression and its cross-ancestry extensions
- June 2024 **2024 WNAR/IMS/Graybill, Fort Collins, CO**  
 Generalized heterogeneous functional model with applications to mobile health data
- May 2024 **Purdue-Lilly Statistics Seminar, Indianapolis, IN**  
 High-dimensional statistical inference for linkage disequilibrium score regression and its cross-ancestry extensions
- May 2024 **STATGEN 2024: Conference on Statistics in Genomics and Genetics, Pittsburgh, PA**  
 High-dimensional statistical inference for linkage disequilibrium score regression and its cross-ancestry extensions
- June 2023 **ICSA Applied Statistics Symposium, Ann Arbor, MI**  
 High-dimensional statistical inference for linkage disequilibrium score regression and its cross-ancestry extensions
- June 2023 **International Purdue Statistics Symposium (IPSS), West Lafayette, IN**  
 Individualized dynamic latent factor model with application in mobile health data
- May 2023 **Midwest Machine Learning Symposium (MMLS), Chicago, IL**  
 Deep generative modeling for nonignorable missing data with nonparametric identifiability
- May 2023 **BFF8, Cincinnati, OH**  
 Statistical inference for high-dimensional block-wise missing data
- Dec. 2022 **CMStatistics (Hybrid), King's College London, UK**  
 An Empirical Bayes Regression for Multi-tissue eQTL Data Analysis
- Nov. 2022 **Lecture Series-Trends in Statistics, National University of Singapore, Singapore**  
 Statistical inference for high-dimensional block-wise missing data
- Apr. 2022 **CCBB Seminar, Indiana University School of Medicine, Indianapolis, IN**  
 Statistical inference for high-dimensional block-wise missing data
- Apr. 2022 **New Challenges and Novel Solutions in Statistics and Data Science, Irvine, CA**  
 An Empirical Bayes Regression for Multi-tissue eQTL Data Analysis

- March 2022 **ENAR Spring Meeting, Houston, TX**  
 Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort
- March 2022 **Seminar, Rutgers University, New Brunswick, NJ**  
 Statistical inference for high-dimensional block-wise missing data
- Feb. 2022 **Biostatistics Seminar, University of Louisville, Louisville, KY**  
 Statistical inference for high-dimensional block-wise missing data
- Feb. 2022 **Statistics Seminar, Indiana University-Purdue University Indianapolis, Indianapolis, IN**  
 Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort
- Jan. 2022 **Bioinformatics Seminar, Purdue University, West Lafayette, IN**  
 Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort
- Nov. 2021 **EAC-ISBA (Hybrid), Atlantic City, New Jersey, USA and Kunming, Yunnan, China**  
 An empirical Bayes regression for multi-tissue eQTL data analysis
- Oct. 2021 **NESS (Hybrid), Kingston, RI, USA**  
 Statistical inference for high-dimensional linear regression with blockwise missing data
- Sep. 2021 **The ICSA Applied Statistics Symposium, Online**  
 Statistical inference for high-dimensional linear regression with blockwise missing data
- Aug. 2021 **JSM, Online**  
 Integrating multi-source block-wise missing data in model selection
- March 2021 **ENAR Spring Meeting, Online**  
 Heterogeneous mediation analysis for causal inference
- Dec. 2020 **The ICSA Applied Statistics Symposium, Online**  
 Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data
- Nov. 2020 **Queen's University, Ontario, Canada**  
 Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data
- Dec. 2019 **Fudan University, Shanghai, China**  
 Integrating multi-source block-wise missing data in model selection
- Dec. 2019 **East China Normal University, Shanghai, China**  
 Integrating multi-source block-wise missing data in model selection
- Dec. 2019 **The 11th ICSA International Conference, Hangzhou, China**  
 Integrating multi-source block-wise missing data in model selection
- Nov. 2019 **Workshop on LEADERS, New Haven, CT, USA**  
 Heterogeneous mediator selection for high-dimensional data
- June 2019 **The ICSA Applied Statistics Symposium, Raleigh, NC, USA**  
 Integrating multi-source block-wise missing data in model selection
- Oct. 2017 **Yunnan University, Kunming, China**  
 Variable selection for highly correlated predictors

[Contributed](#)

- Aug. 2024 **JSM, Portland, OR**  
Individualized dynamic latent factor model with application in mobile health data
- Aug. 2024 **New Researchers Conference (NRC), Corvallis, OR**  
Generalized heterogeneous functional model with applications to mobile health data
- Aug. 2022 **JSM, Washington, DC, USA**  
An empirical Bayes regression for multi-tissue eQTL analysis
- Aug. 2020 **JSM, Online**  
Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data
- Aug. 2018 **JSM, Vancouver, Canada**  
Variable selection for highly correlated predictors
- Nov. 2018 **Bohrer Workshop, Champaign, IL, USA**  
Integrating multi-source block-wise missing data in model selection
- Nov. 2017 **Bohrer Workshop, Champaign, IL, USA**  
Variable selection for highly correlated predictors

---

## Consulting Experience

- 2016–2019 **Statistical Consultant, Illinois Statistics Office, UIUC**  
Provided statistical consultation to 40+ clients
- Clients from Eli Lilly and Company, Pace Suburban Bus, Accel Entertainment, BioFortis Inc, Carle Foundation Hospital, and University of Illinois
  - In the fields of biology, business, ecology, engineering, finance, marketing, medicine, psychology, sociology, etc.
- Covered a wide range of services
- Grant proposal and medical protocol review
  - Experimental design (survey design, clinical trial design)
  - Data analysis (ANOVA, regression, factor analysis, mixed models, etc.)
  - Prediction (market prediction, yield prediction)
  - Software tutorial and technical report writing

---

## Referee Experience

- Journal of the American Statistical Association (17)  
Journal of the Royal Statistical Society: Series B (5)  
Annals of Statistics (1)  
Journal of Machine Learning Research (2)  
Electronic Journal of Statistics (3)  
Journal of Nonparametric Statistics (2)  
Journal of Computational and Graphical Statistics (2)  
Technometrics (1)  
PLOS One (1)  
PLOS Genetics (1)  
Biometrics (5)  
Bioinformatics (1)

The Canadian Journal of Statistics (2)  
Statistics in Medicine (6)  
Statistica Sinica (6)  
Lifetime Data Analysis (1)  
Science China Mathematics (1)  
Journal of Multivariate Analysis (1)  
TEST (1)  
Stat (1)  
Computational Statistics and Data Analysis (2)  
Fertility and Sterility (1)  
Biostatistics (2)

## Service

- 2025–2026 **Chair of Colloquium Committee**  
Purdue University
- December 2024 **External Reviewer of Natural Sciences and Engineering Research Council of Canada (NSERC)**  
Natural Sciences and Engineering Research Council of Canada (NSERC)
- August 2024 **Session Organizer of Novel Statistical Methods for Mobile and Wearable Device Data**  
JSM, Portland, OR, USA
- August 2024 **Session Chair of New Statistical Methods for High-dimensional Inference**  
JSM, Portland, OR, USA
- August 2024 **Chair of Funding Panel**  
New Researchers Conference (NRC), Corvallis, OR
- March 2024 **NSF Panelist of Statistics Program**  
National Science Foundation
- 2024–2025 **Member of the 9th Bayesian, Fiducial, and Frequentist (BFF9) Conference Committee**  
Purdue University
- 2024–current **Member of Graduate Program Committee**  
Purdue University
- 2023–current **Member of Talent-Based Hires Committee**  
Purdue University
- 2021–current **Chief Executive Officer of Purdue Statistics Alumni Association (PSAA)**  
Purdue University
- 2021–2025 **Member of SLDS Student Paper Award Committee**  
American Statistical Association
- 2022–2023 **Member of Academic Program Committee for International Purdue Statistics Symposium (IPSS)**  
Purdue University

- 2022–2023 **Member of Local Organizing Committee for International Purdue Statistics Symposium (IPSS)**  
Purdue University
- June 2023 **Session Organizer of Purdue Statistics Alumni**  
IPSS, West Lafayette, IN
- June 2023 **Session Organizer of Recent Developments in Data Integration**  
IPSS, West Lafayette, IN
- March 2023 **Panel Organizer of Networking and career development**  
PSAA March Event, online
- March 2023 **Organizer of Purdue Statistics Alumni Association (PSAA) March Event**
- December 2022 **Session Organizer of Recent Developments in Statistical Machine Learning**  
CMStatistics, London, UK
- August 2022 **Session Organizer of Recent Developments in Causal Inference**  
JSM, Washington, DC, USA
- March 2022 **Panel Organizer of Collaboration and Networking**  
PSAA March Event, online
- March 2022 **Organizer of Purdue Statistics Alumni Association (PSAA) March Event**
- 2022 **Member of External Review Committee**  
Purdue University
- 2021–2022 **Member of Data Science Search Committee**  
Purdue University
- 2017–2019 **Committee Member of Graduate Student Invited Seminar**  
The University of Illinois at Urbana–Champaign
- July 2017 **Session Chair of Machine Learning Applications**  
JSM, Baltimore, MD, USA
- June 2017 **Session Chair of Prediction and inference for high-dimensional data and non-parametric models**  
The ICSA Applied Statistics Symposium, Chicago, IL, USA

---

## Professional Memberships

- American Statistical Association  
Institute of Mathematical Statistics  
International Chinese Statistical Association